PERGAMON Computers and Mathematics with Applications 38 (1999) 35–50

An International Journal
computers & mathematics
with applications

www.elsevier.nl/locate/camwa

A Numerical Comparison of Seven Grids for Polynomial Interpolation on the Interval

J. P. BOYD
Department of Atmospheric, Oceanic, and Space Science
2455 Hayward Avenue, Ann Arbor, MI 48109, U.S.A.
jpboyd@engin.umich.edu

(Received May 1998; revised and accepted Febuary 1999)

Abstract—Seven types of Chebyshev-like grids in one dimension are compared according to four different criteria for accuracy. The grid which minimizes the Lebesgue constant is the best because it performs fairly well by all four criteria. For the same reason, minimizing the Lebesgue constant seems to be the most useful measure of optimality because grids that are good according to this criterion are good when measured by other criteria, too. CFM-optimality, which is the property that all cardinal functions (Lagrange fundamental polynomials) have maxima at the interpolation points, seems to be the least-discriminating criterion because all seven grids generate cardinal functions that have maxima at or very near the interpolation points. The difference between these grids on all four criteria, always less than a factor of two and usually much smaller, are sufficiently modest so that the final choice between grids should probably not be made because of accuracy, but rather based on other criteria such as ease-of-programming, analytical simplicity, conformality with other approximations, and timestep in applications to partial differential equations. To explore such nonaccuracy issues, we also compared six of the grids on the basis of maximum allowed timestep when the grid is used to discretize the spatial coordinate and an explicit scheme is used for time-marching. The Lebesgue-optimal grid is also nearly optimal in the sense of allowing the longest timestep. With a pseudospectral (collocation) algorithm, the Lebesgue grid allows a timestep three-halves as long as that of the Legendre-Lobatto spectral element method.

Keywords—Lagrange interpolation, Fekete grid, Lebesgue constants, Spectral element.

1. INTRODUCTION

What choice of grid points is best for interpolation on the interval? The triangle? The tetrahedron? Although the mechanics of interpolation on the interval was well understood by Lagrange more than two centuries ago, the controversy about which points are optimum still rages. Chen and Babuška [1–3] recommend the Lebesgue-optimal points for the interval [4], but retreat to Erdős-optimal grids for the triangle and tetrahedron, which are easier to compute. In contrast, Bos [5] and Wingate and Taylor [6–8] advocate Fekete-optimal (VDM-optimal) grids for the triangle. To exploit the computational efficiency of tensor product transformations, Sherwin and

This work was supported by NSF grant OCE9521133. I thank B. A. Wingate and M. A. Taylor for showing me their work on triangles in advance of publication, which inspired this study of interpolation on the interval. I also thank the reviewers for helpful comments.

0898-1221/99/$ - see front matter. Typeset by AMS-TEX
PII: S0898-1221(99)00204-7

Karniadakis [9–12] and Wingate and Boyd [13] use a different, noninterpolating algorithm for the triangle with roughly twice as many grid points as polynomial degrees-of-freedom. Királyfalvi and Szabó [14] use Erdős-optimal grids for both the quadrilateral and the triangle so that elements of different geometric type can be mixed conformally. Hesthaven [15] has compared several grids on the interval, but compared them only by a single criterion of optimality.

In this note, we pursue the modest goal of comparing various grids on the interval according to these multiple standards of "optimality". At the end, we shall make a few comments about the implications of our findings for the triangle and tetrahedron.

The quest for an optimal set of grid points for polynomial interpolation on the interval has been greatly complicated by the fact that there are many different measures of "optimal". One of these is defined in terms of the "Cauchy error factor"

$$\omega(x) \equiv \prod_{k=1}^{N} (x - x_k)\,. \tag{1}$$

Cauchy proved that, denoting the N-point interpolating polynomial of degree $(N-1)$ by $L_{N-1}(x)$,

$$f(x) - L_{N-1}(x) = \frac{f^{(N)}(\xi)}{N!}\omega(x), \tag{2}$$

where f is the function being interpolated, $f^{(N)}$ is its N^{th} derivative, and ξ is some point on the interval spanned by the grid points and x. Since the only part of this error which can be manipulated by the choice of grid points is $\omega(x)$, the goal of "Cauchy optimality" is to minimize ω.

The other measures depend upon the set of polynomials $C_j(x)$ which are one at the j^{th} grid point and zero at all the others, known variously as the "Lagrange polynomials", "fundamental polynomials of Lagrange interpolation", "cardinal basis", or the term that we shall use, the "cardinal functions":

$$C_j(x) = \frac{\prod_{k=1,\,k\neq j}^{N} (x - x_k)}{\prod_{k=1,\,k\neq j}^{N} (x_j - x_k)} = \frac{\omega(x)}{(x - x_j)\omega'(x_j)}, \tag{3}$$

where ω' is the first derivative of ω with respect to x.

In functional analysis, the optimal grid is that which minimizes the so-called Lebesgue constant defined below. Because the "Lebesgue optimal" grid is difficult both numerically and theoretically, two related criteria, Erdős optimality and Fejér optimality, have been widely used as proxies or substitutes. Because of their close relationship with the Lebesgue criterion, we shall define Erdős and Fejér criteria below, but not otherwise discuss them.

The cardinal function $C_j(x)$ decays rapidly as $|x - x_j|$ increases. Yet another measure of optimality is that each cardinal function should be a maximum at $x = x_j$, which will be dubbed CFM optimality ("Cardinal-Function-Maximizing").

In terms of a fixed spatial basis $\{\phi_j(x)\}$, $j = 1, \ldots, N$, the computation of the interpolant of an arbitrary function requires the inversion of the generalized Vandermonde matrix defined by

$$\mathcal{VDM}_{ij} = \phi_j(x_i), \qquad i = 1, 2, \ldots, N, \quad j = 1, 2, \ldots, N, \tag{4}$$

Fekete [16] argued that a grid which maximized the determinant of this matrix (for a fixed ϕ_j) would be nearly optimal. Such Vandermonde-maximizing grids are often called "Fekete" grids, but this term is sometimes more loosely applied to grids that are minima of electrostatic repulsion of point charges at the grid point [6]. Bos [5] and Chen and Babuška [1–3] simply refer to this grid with the subscript "VDM", so we shall refer to this grid as VDM-optimal. Bos [5] has advocated VDM-optimal grids for simplexes in one or more dimensions; he showed that the grids

can be computed by choosing a fixed basis and then maximizing the corresponding Vandermonde determinant as the grid points are varied, which is a fairly straightforward numerical procedure.

In this article, we shall concentrate on these most fundamental measures of optimality: Cauchy, Lebesgue, CFM, and VDM. For completeness, we list six definitions below.

DEFINITION 1.1 CAUCHY OPTIMALITY. Minimize

$$\Omega_N \equiv 2^{N-1} \max_{x\in[-1,1]} |\omega(x)|, \tag{5}$$

where 2^{N-1} is a scaling factor chosen so that $\Omega_N \equiv 1$ for all N for the Chebyshev-interior grid, which is the best possible choice of points for this optimality criterion.

DEFINITION 1.2 LEBESGUE OPTIMALITY. Minimize the Lebesgue constant

$$\Lambda_N \equiv \max_{x\in[-1,1]} \sum_{j=1}^{N} |C_j(x)|. \tag{6}$$

DEFINITION 1.3 CFM ("CARDINAL-FUNCTION-MINIMIZING") OPTIMALITY. Minimize

$$\mathrm{CFM}_N \equiv \max_j \left\{ \max_{x\in[-1,1]} |C_j(x)| \right\}. \tag{7}$$

DEFINITION 1.4 VDM ("VANDERMONDE-DETERMINANT-MAXIMIZING") OPTIMALITY. Maximize

$$\mathcal{V}_N \equiv \text{DETERMINANT (VDM)} \tag{8}$$

for a fixed basis $\{\phi_j\}$ but varying grid point $\{x_i\}$ where $\mathcal{VDM}$ is the Vandermonde matrix defined by equation (4) above.

DEFINITION 1.5 FEJÉR OPTIMALITY. Minimize the L_∞ norm of the sum of the squares of the cardinal functions:

$$F_N \equiv \max_{x\in[-1,1]} \sum_{j=1}^{N} (C_j(x))^2. \tag{9}$$

DEFINITION 1.6 ERDŐS OPTIMALITY. Minimize the L_2 norm of the integral of the squares of the cardinal functions:

$$E_N \equiv \int_{-1}^{1} \sum_{j=1}^{N} (C_j(x))^2. \tag{10}$$

In this article, we compute the first four of these optimality measures for seven representative sets of grid points. The "Chebyshev-interior" grid consists of the roots of the Chebyshev polynomial $T_N(x)$, which all lie on the interior of the interval $x \in [-1, 1]$; Chebyshev proved that this is the Cauchy-optimal grid.

The second, third, and fourth grids are Lobatto grids for Gegenbauer polynomials. The "Lobatto" points are those which appear in a Gauss-Lobatto quadrature formula and consist of the endpoints $x = \pm 1$ plus the interior extrema of the Gegenbauer polynomial of degree $(N - 1)$. The Gegenbauer polynomials are the set of polynomials orthogonal with respect to the weight function $w(x) \equiv (1 - x^2)^\alpha$. The "Chebyshev-Lobatto" grid has $\alpha = -1/2$, which yields the usual Chebyshev polynomials. The "Legendre-Lobatto" grid has $\alpha = 0$, which gives the Legendre polynomials. Lastly, we also tried $\alpha = 1/2$, which gives the Lobatto grid for Chebyshev polynomials of the second kind. Empirically, this grid seems to be closest, among the family of Gegenbauer-Lobatto grids, to the set of points which are Lebesgue-optimal.

Chen and Babuška [1–3] numerically computed the "Lebesgue-optimal" and "Erdős-optimal" grids for up to twenty points. These will be our fifth and sixth sets of grid points.

Lastly, we also try the "expanded-Chebyshev" grid proposed by Brutman [17], which is the Chebyshev-interior grid stretched by a linear mapping so that the endpoints are moved to ± 1:

$$x_k(\text{expanded-Chebyshev}) \equiv \frac{\cos\left((2k-1)\pi/(2N)\right)}{\cos(\pi/(2N))}, \qquad k = 1, 2, \ldots, N. \tag{11}$$

Empirically, this is a very good approximation to the Lebesgue-optimal grid points, which are not known analytically.

The "Legendre-Lobatto" grid is CFM-optimal, VDM-optimal, and Fejér-optimal [18]. The j^{th} cardinal function can be expressed in terms of the Vandermonde determinant as [6]

$$C_j(x) = \frac{\text{determinant}\left(\text{VDM}(x_1, x_2, \ldots, x_j = x, \ldots, x_N)\right)}{\text{determinant}\left(\text{VDM}(x_1, x_2, \ldots, x_j, \ldots, x_N)\right)}, \tag{12}$$

where the notation means that the x_j is replaced by the coordinate x in the Vandermonde matrix in the numerator. From this, it follows that a VDM-optimal grid is also CFM-optimal in the sense that the j^{th} cardinal function has its maximum of one at x_j for each j. (This property generalizes from the interval to approximation in squares, triangles, and general regions in higher dimensions.) However, CFM-optimality does not necessarily imply VDM-optimality because there could be multiple sets of points which are all Cardinal-Function-Minimizing, but have different Vandermonde determinants. On the interval, however, [18] proved that the VDM-optimal grid is the Legendre-Lobatto grid. Bos (to be published) has recently shown that a tensor product Legendre-Lobatto grid on the square is VDM-optimal. For grids on a triangular region in a two-dimensional space, however, there are multiple CFM-optimal grids that are not all VDM-optimal.

The most obvious choice of grid is evenly-spaced, but it has been known since the turn-of-the-century that this is a horrible choice. The Lebesgue constants rise exponentially rather than logarithmically with N and the interpolation generally diverges [19,20].

The seven grids that we shall consider are all "Chebyshev-like" in the sense that all have grid points clustered near the ends of the interval at $x = \pm 1$ with a nearest-neighbor separation near the boundaries proportional to $1/N^2$, rather than to $1/N$ as for an evenly spaced grid. It is the subtle differences in performance within the family of "Chebyshev-like" grids that are the source of the continuing controversy, not the huge difference between a Chebyshev grid and an evenly-spaced grid.

Unfortunately, Chebyshev-like grids, when applied to spatial coordinates of a time-dependent partial differential equation, lead to "stiff" systems of ordinary differential equations in time. Another measure of "optimal" is: Which grid allows the longest timestep? In Section 6, we compare grids according to timestep optimality.

2. CAUCHY OPTIMALITY

Figure 1 compares the Cauchy optimality, defined by equation (5) as the L_∞ norm of the error factor $\omega(x) = \prod_{k=1}^{N}(x - x_k)$, for the seven grids. The Chebyshev-interior grid, that is, the roots of $T_N(x)$, is the best with $\Omega_N \equiv 1$ for all N. The Chebyshev-Lobatto grid is twice as bad with Ω_N asymptoting to 2 as $N \to \infty$. The Legendre-Lobatto (Fekete) grid is not too good either, with its Cauchy optimality asymptoting to roughly $\sqrt{2}$. The Erdos-optimal grid is somewhat better, the Chebshev-Lobatto-of-the-second-kind is better still.

Strikingly, the expanded-Chebyshev grid and the Lebesgue-optimal grid, which are almost superimposed on top of one another on the graph, seem to be asymptoting to one from above. Even for $N = 10$, the Cauchy optimality for these grids is no worse than 1.15. The conclusion is that Cauchy-optimality and Lebesgue-optimality are not mutually incompatible. Instead, the grid which is optimal in terms of miniming the Lebesgue constant (thick solid curve) is *almost* optimal

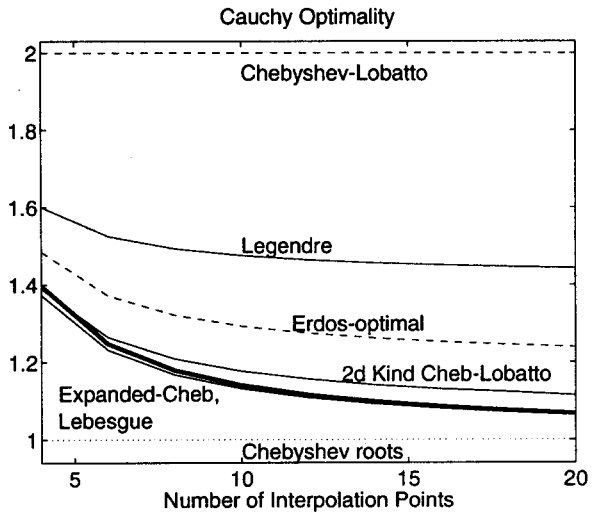

Figure 1. The Cauchy optimality function $\Omega_N \equiv ||\omega(x)||_\infty$ for various grids and different numbers N of interpolation points. The thick curve is the Lebesgue-optimal curve; it almost indistinguishable from the expanded-Chebyshev grid. The thin dotted curve at the bottom is the Chebyshev-interior grid, for which $\Omega_N = 1$ for all N.

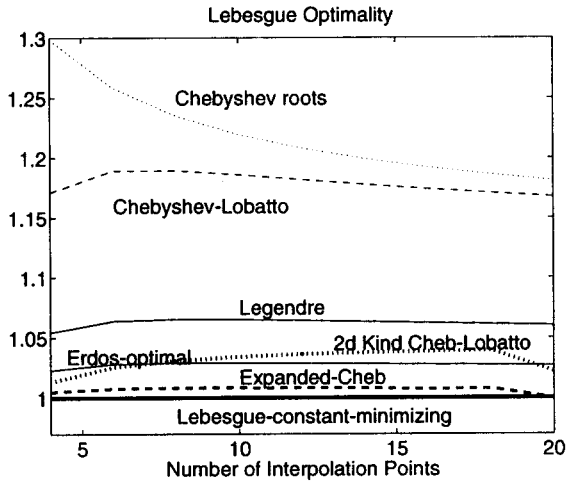

Figure 2. Lebesgue optimality, divided by the smallest possible Lebesgue constant for a given N, for seven different grids. The scaled Lebesgue optimality is identically equal to one (thick horizontal line) for the Lebesgue-optimal grid; the unscaled Lebesgue constants (not shown) grow proportional to $(2/\pi)\log(N)$ [19]. The Legendre grid is also the VDM-optimal grid and the Fekete grid.

in terms of minimizing the error in Cauchy's Theorem, too. The expanded-Chebyshev grid has the same virtues combined with the added blessing of an analytic formula for the gridpoints.

3. LEBESGUE OPTIMALITY

Brutman [19] has dubbed Lebesgue-optimality the "functional analysis" criterion because the Lebesgue constant is the L_∞ norm of the interpolation operator. The two standard Chebyshev grids, so widely used in numerical analysis, are mediocre in the sense that their Lebesgue constants are about 25% larger than the best possible. The VDM-optimal grid, however, is only about 5% above the optimum. The expanded-Chebyshev grid is barely 1% above the values of the true Lebesgue-optimal grid, whose points can only be found numerically by a nonlinear programming method [4]; Brutman [17] proves that the scaled Lebesgue constants for the expanded-Chebyshev grid can exceed one by no more than $0.127/\log(N)$ for moderate and large N.

It follows that in terms of Lebesgue constants, one pays only a slight price by using the Legendre-Lobatto (Fekete/VDM-optimal) or Chebyshev-of-the-second-kind-Lobatto. The expanded-Chebyshev grid is so good that it hardly seems worth the bother of copying the true Lebesgue-optimal grid points from Chen and Babuška [3].

4. CARDINAL-FUNCTION-MINIMIZING (CFM) OPTIMALITY

Figure 3 compares the CFM optimality for six of our seven grids. The Chebyshev-interior grid, which is the only grid that does not include $x = \pm 1$, is not shown because the cardinal functions associated with the two points nearest the endpoints rise to approximately 1.27, roughly independent of N, at $x = \pm 1$. However, the maxima of all the other cardinal functions for all N are 1.04 or smaller, comparable with the CFM-optimality of the other six grids.

The most remarkable feature of the graph is how little the CFM-optimality varies from one grid to another. The Chebyshev-of-the-second-kind-Lobatto grid, which is the worst of the six illustrated, still has cardinal function maxima no worse than 1.033 versus the best possible of 1. The Lebesgue-optimal grid has cardinal function maxima no larger than 1.015.

CFM optimality is not a very discriminating criterion. Six of the seven grids tested here are *very* close to optimal. The one that is not, the Chebyshev-interior grid, is the best in terms of Cauchy-optimality and has been successfully used in literally hundreds of papers over the last half-century. The fact that two of its cardinal functions rise steeply towards $x = \pm 1$ seems to be quite irrelevant to the good performance of this set of grid points.

Still, the CFM-optimal grid, alias the Fekete grid, has been used successfully in a wide range of spectral element calculations because it is also the Legendre-Lobatto grid, which arises naturally in a variational formulation of boundary value problems. We have already seen in the previous two sections that the Legendre-Lobatto grid gives Lebesgue constants no more than 5% worse than the best possible. It is significantly inferior to other grids in terms of Cauchy-optimality with the second worst score $\Omega_N \approx 1.41$. However, the rapid convergence of spectral element approximations with increasing N implies that this difference would be hard to observe in head-to-head comparisons with other grids for solving differential equations.

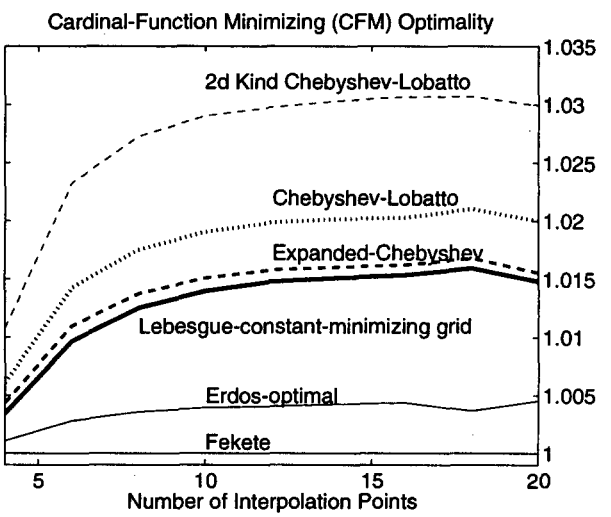

Figure 3. The CFM optimality (maximum of the cardinal functions) for various grids and N. The thick solid curve is the Lebesgue-optimal grid. The CFM-optimal grid, which is the Legendre-Lobatto grid, always has a CFM-optimality of one.

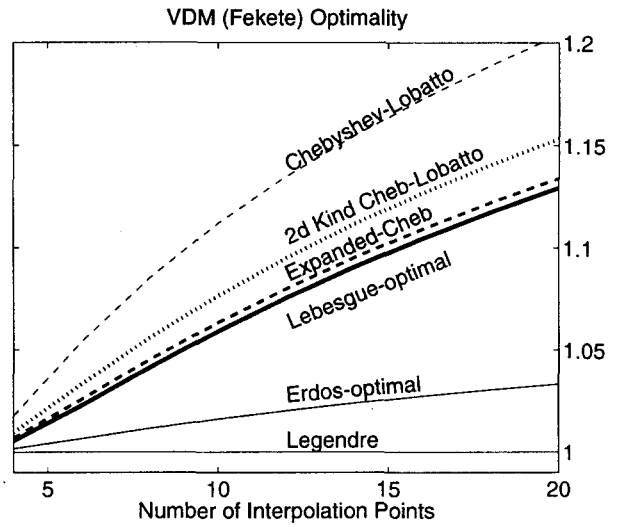

Figure 4. VDM (Fekete) optimality. The plotted quantity is the inverse of the absolute value value of the Vandermonde determinant after scaling by the determinant for the Legendre-Lobatto grid. (The reciprocal of $\det(\mathcal{VDM})$ is graphed, instead of the determinant itself, so that the higher the curve on the plot, the worse the performance of the grid, as for the previous three graphs.) After being scaled by itself, the Legendre-Lobatto grid always has a reciprocal-VDM optimality of one. The CFM-optimality for the Chebyshev-roots grid is not shown because it ranges from 1.6 to 2.04 for the plotted range of interpolation points, well above the upper limit of the axis.

5. VANDERMONDE-MATRIX-DETERMINANT-MAXIMIZING (VDM) OPTIMALITY

Figure 4 shows that the relative performance of the grids according to VDM (alias, Fekete) optimality is very similar to CFM-optimality except that the roles of the Chebyshev-roots and Chebyshev-Lobatto grids, which are the worst pair in both cases, have changed places. The similarity of Figures 3 and 4 is not surprising since the same grid, the Legendre-Lobatto points, are optimal according to the both CFM and VDM criteria.

The rise in $\det(\mathcal{VDM})$ with N for the Lebesgue grid, which is preferred by functional analysis, suggests that VDM-optimality is not merely indiscriminate, but actually misleading.

6. TIMESTEP OPTIMALITY

The length of the maximum stable timestep is an important issue in solving time-dependent partial differential equations. Unfortunately, for polynomial pseudospectral methods, the eigenvalues of the derivatives matrices have largest eigenvalues which are $O(N^2)$ for the first derivative and $O(N^4)$ for the second derivative, versus $O(N)$ and $O(N^2)$ for a finite difference method of low, fixed order on an evenly-spaced grid. For discretizing the spatial dependence of time-dependent partial differential equations, these large, spurious eigenvalues of the differentiation matrices imply that spectral and spectral element methods are notoriously "stiff", that is, require a short timestep for an explicit time-marching algorithm, or alternatively, the additional complications and expense of an implicit time-integration scheme.

This motivates another definition of optimality.

DEFINITION 2. *The "timestep optimality" of a grid is inversely proportional to the largest magnitude of the eigenvalues of its differentiation matrix.*

The diffusion equation in one dimension with homogeneous Dirichlet boundary conditions is a simple example:

$$u_t = u_{xx}, \qquad u(-1) = u(1) = 0, \tag{13}$$

where the subscripts denote differentiation with respect to the coordinate. For any grid which includes the endpoints ± 1, spectral discretizations approximate the solution as

$$u(x,t) = \sum_{j=2}^{N-1} u_j(t) C_j(x), \tag{14}$$

where the boundary conditions are imposed by simply omitting the cardinal functions corresponding to the endpoints. (The pseudospectral discretization using the Chebyshev-interior grid is slightly more complicated because the boundary conditions have to be imposed separately, so in this section, we shall omit this grid.)

Two different discretizations are popular. The "pseudospectral" or "collocation" method demands that the residual of the differential equation should be zero at each of the $(N-2)$ interior points, which gives

$$\vec{u}_t = \vec{D}^{(\text{pseudo})} \vec{u}, \tag{15}$$

where $\vec{u}$ is a column vector containing the time-dependent values of u at the interior spatial grid points, $u_j(t)$, and $\vec{D}$ is an $(N-2) \times (N-2)$ square matrix whose elements are

$$D_{ij}^{(\text{pseudo})} \equiv \frac{d^2 C_{j+1}}{dx^2}(x_{i+1}), \qquad i = 1, \ \ldots, (N-2), \quad j = 1, \ldots, (N-2). \tag{16}$$

The alternative discretization, which is variously called the "h-p/finite element" or "spectral element" method, employs a variational formalism. The integral of the residual function, multiplied by a set of test functions, must be equal to zero. The integral of the test function with the second derivative is always integrated-by-parts to be become the negative of the integral of the first derivative of the test function with first derivative of the solution. In the early days of finite elements, it was customary to perform the integrals exactly, but this has the unpleasant consequence that the matrix multiplying the time derivative is no longer the identity matrix, as for collocation, but a matrix that must be factored and then backsolved at every timestep. Today it is usual to sidestep this problem by

(i) employing a cardinal function basis and

(ii) replacing the exact integral by an N-point quadrature of the form

$$\int_{-1}^{1} f(x)\, dx \approx \sum_{j=1}^{N} w_j f(x_j), \tag{17}$$

where the quadrature weights are given by

$$w_j \equiv \int_{-1}^{1} C_j(x)\, dx, \tag{18}$$

so that one is simply approximating the integrand by its interpolant in the form of a cardinal function series and then integrating term-by-term. The "variational" approximation becomes

$$\vec{u}_t = \vec{D}^{(\text{stiff})} \vec{u}, \tag{19}$$

where the elements of the "stiffness" matrix are

$$D_{ij}^{(\text{stiff})} \equiv -\frac{1}{w_i} \sum_{k=1}^{N} w_k \frac{dC_{i+1}}{dx}(x_k) \frac{dC_{j+1}}{dx}(x_k), \quad i = 1, \ldots (N-2), \quad j = 1, \ldots (N-2). \tag{20}$$

(Actually, the matrix on the right-hand side of the diffusion equation is the usual stiffness matrix divided by the inverse of the diagonal mass matrix, but this minor abuse of the usual finite element terminology is convenient because what we call the "stiffness matrix" is what must be compared with the pseudospectral matrix.)

For the Legendre-Lobatto (Fekete) grid, the stiffness and pseudospectral matrices are *identical*. For the other grids, these matrices are different, and there are thus, two slightly different measures of timestep optimality for each grid, depending on which spatial discretization we choose.

For either the pseudospectral or spectral element formalism, if the initial condition is proportional to an eigenvector of $\vec{D}$, then the system of ordinary differential equations in time collapses to the single differential equation

$$\frac{da}{dt} = \lambda a, \tag{21}$$

where $a(t)$ is the amplitude of the eigenfunction and λ is the eigenvalue. An explicit time-marching scheme for equation (21) will become unstable when the timestep Δt satisfies the inequality

$$\Delta t > \frac{\text{constant}}{\lambda}, \tag{22}$$

where the constant is an $O(1)$ number that depends only the algorithm (third-order Runge-Kutta versus second-order Adams-Bashforth, etc.). The maximum timestep for the *system* of ordinary differential equations in time is inversely proportional to the *largest* eigenvalue (in magnitude) of the differential equation.

All this is standard stability theory; with apologies for glossing over some technicalities, we refer the reader to [20–22] for a more complete treatment. The important point is that timestep optimality is inversely proportional to the eigenvalues of the differentiation matrix. For more complicated problems, this matrix will a combination of first and second derivatives with partial derivatives, etc. However, the basic concept can be adequately illustrated by simply computing the eigenvalues of the second derivative matrix $\vec{D}$ for various grids. Canuto *et al.* [22, p. 100] show that the largest Chebyshev-Lobatto and Legendre-Lobatto eigenvalues are asymptotically proportional to $0.047N^4$ and $0.026N^4$, but results for other grids have not been previously published.

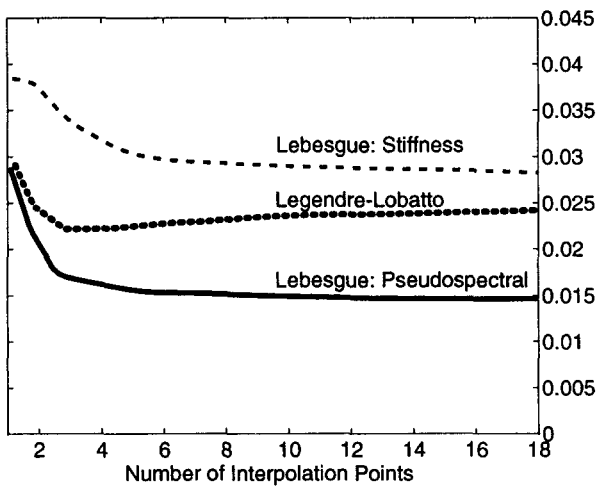

Figure 5. Maximum eigenvalues of the pseudospectral and stiffness matrices with Dirichlet boundary conditions for the Lebesgue-optimal and Legendre-Lobatto (Fekete-optimal) grids, all scaled by dividing by the fourth power of the number of interpolation points. *Smaller* is *better* in the sense of allowing a longer stable timestep for explicit time-integration. The Legendre-Lobatto stiffness and pseudospectral matrices have identical elements for this case, so only a single curve is shown for this grid.

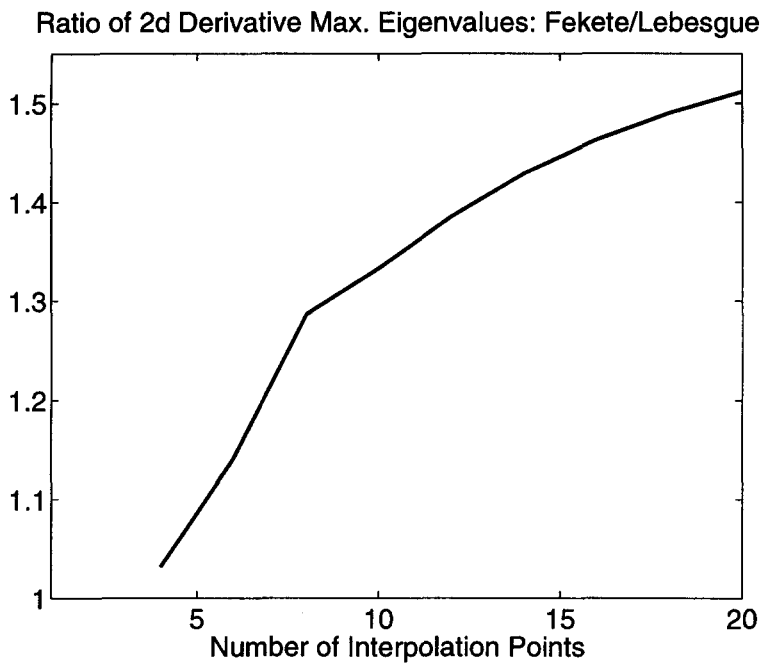

Figure 6. Ratio of the maximum eigenvalue of the stiffness or pseudospectral matrix for the Legendre-Lobatto grid, which is CFM-optimal, Fekete-optimal, and VDM-optimal to that for the pseudospectral matrix of the second derivative on the Lebesgue-optimal grid.

Figure 5 shows the largest eigenvalues for the Lebesgue-optimal and Legendre-Lobatto grids. We have restricted the graph to just these two sets of points because

(i) the curves for the expanded-Chebshev grid are almost indistinguishable from the Lebesgue grid,

(ii) the curves for the Chebyshev-of-the-second-kind-Lobatto are close to the Lebesgue curves, too, and

(iii) the Erdos grid is intermediate between the Lebesgue and Legendre-Lobatto in grid point locations and also in timestep optimality.

The Lobatto grid for the ordinary Chebyshev polynomials is quite bad as already found by Canuto *et al.* [22].

The clear winner is the pseudospectral Lebesgue-optimal discretization. Figure 6 shows the ratio of the maximum eigenvalue for the VDM-optimal/Legendre-Lobatto grid to the Lebesgue-optimal grid. For small N, there is little difference. When 20 grid points are used, however, the Lebesgue-optimal grid allows a maximum stable timestep half again as large as for the standard Legendre-Lobatto grid—and it is superior in both Cauchy-optimality and Lebesgue-optimality!

Remarkably, however, the stiffness matrix for the Lebesgue-grid imposes a timestep limit noticeably more severe than for the Legendre-Lobatto grid. Perhaps this is not too surprising in the sense that the Legendre-Lobatto grid gives a Gaussian quadrature formula for integration, and thus, does especially well in the variational formalism, which employs integration and integration-by-parts. Why, however, does the Lebesgue grid (in the pseudospectral discretization) do so well compared to other grids?

Heuristically, the stability limit for both wave equations and diffusion equations is that the maximum stable timestep is proportional to and of the same magnitude as the length of time required to propagate or diffuse from one grid point to the next. This is unfortunately a very short time when the grid points are clustered densely near ± 1, as true of all the grids here. The Lebesgue grid allows a longer timestep because the clustering of points near the boundary is not as extreme as for the Legendre-Lobatto grid.

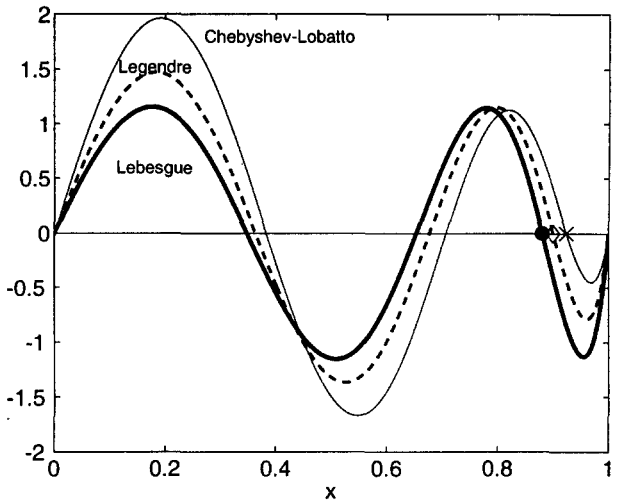

Figure 7. The Cauchy factor $\omega(x)$ for three grids. Thin solid line: Chebyshev-Lobatto grid (ordinary Chebyshev polynomials). Dashed curve: Legendre-Lobatto (VDM-optimal; Fekete grid). Thick solid line: Lebesgue-optimal. The grid points nearest the endpoint, which control the timestep limit, are marked by symbols near the right edge of the graph. Disk: Lebesgue. Diamond: Legendre-Lobatto. X: Chebyshev-Lobatto.

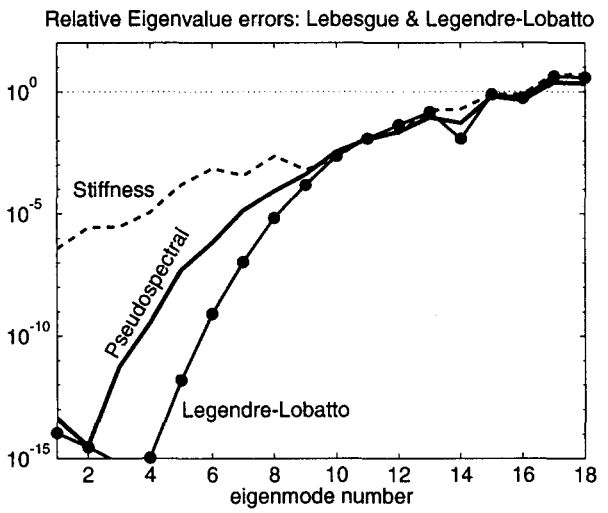

Figure 8. Relative errors in the eigenvalues of the second derivative operator with Dirichlet boundary conditions as approximated by the Legendre-Lobatto grid [solid curve with disks] (either pseudospectral or spectral element, which give the same discretization matrix) and by the Lebesgue grid with a pseudospectral algorithm (solid line) or the spectral element stiffness method (dashed).

Figure 7 compares the Cauchy factor $\omega(x)$ for three different grids. The grid points are the roots of the corresponding ω. The Lebesgue grid is the best of the three in the Cauchy-optimal sense because it varies most uniformly over the interval. It is also the best for timestep-optimality because the grid point nearest $x = 1$ (i. e., the root of $\omega(x)$ nearest $x = 1$, marked by the filled disk) is much farther from the endpoint than the corresponding axis-crossing points for the other grids. The Chebyshev-of-the-second-kind-Lobatto and expanded-Chebyshev ω are not shown because these are almost indistinguishable on the scale of the graph from the Lebesgue-optimal grid, and therefore, share its virtues. It is terribly ironic that the Legendre-Lobatto grids, which are inferior in both Cauchy-optimality and in the maximum allowed timestep, have been used almost exclusively in spectral elements until now [23,24].

The Legendre-Lobatto grid has become so ubiquitous in part because the variational formalism is convenient for proving convergence theorems. Another important issue is: how do these grids compare in accuracy for solving differential equations, as opposed to interpolation? The eigenvalues of the second derivative with Dirichlet boundary conditions on $x \in [-1, 1]$ are given exactly by

$$\lambda_{\text{exact}} = \frac{\pi^2}{4} j^2, \qquad j = 1, 2, \ldots, \tag{23}$$

where j is the mode number. Figure 8 shows the approximation of these eigenvalues by various grids and algorithms. As with any eigenvalue approximation, only about (N/2) to (2 N/3) eigenvalues are accurately approximated by a discretization with N points where the cutoff between "good" and "bad" eigenvalues depends on the user's choice of error tolerance. The Legendre-Lobatto spectral element method is much better than the Lebesgue pseudospectral algorithm for the lowest few modes, and this in turn is much better than the stiffness matrix on the Lebesgue grid. It is remarkable and surprising that small differences in the discretization method and the grid can change the error by as much as $1.E-10$! For moderate mode number (9–12), however, the relative errors are small for all three methods and almost indistinguishable. It is unclear whether the superior accuracy of the Legendre-Lobatto grid for the lowest few eigenvalues, which all the schemes approximate well, would improve overall accuracy for a time-dependent solution of the diffusion equation.

In summary, the pseudospectral method on the Lebesgue grid seems a very attractive alternative to the Legendre-Lobatto spectral element because it allows a significantly longer timestep (50% for moderate and large N). The Lebesgue grid is also better for interpolation, but not as good at approximating eigenvalues of the second derivative. The accuracy penalty, if any, is likely too small to cancel the advantages of the longer timestep allowed by the Lebesgue grid.

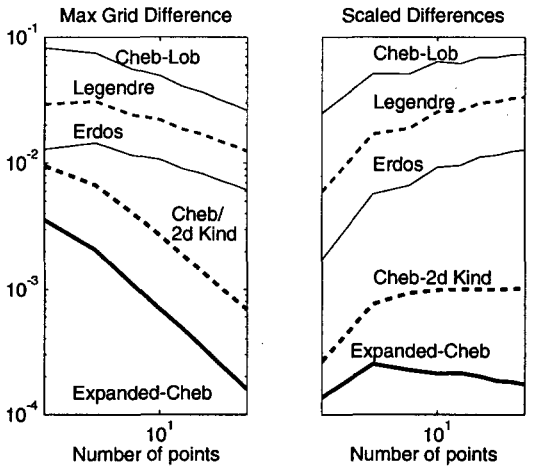

Figure 9. Left panel: Maximum difference in grid points between various grids and the Lebesgue grid, $\max_k |x_k - x_k^{(Lebesgue)}|$, plotted versus the number of interpolation points on a log-log scale. The two grids which are close to the Lebesgue grid are shown as thick curves: the expanded- Chebyshev grid (bottom, closest to Lebesgue) and the Chebyshev-of-the-Second-Kind-Lobatto. Right: the differences scaled by multiplication by N (top three curves) or N^2 (bottom two). The slow growth of the top three curves implies that the Chebyshev-Lobatto, Legendre-Lobatto (CFM-optimal/Fekete-optimal) and Erdős gridpoints are moving closer to their Lebesgue counterparts more slowly than $O(1/N)$. In contrast, the two "almost-Lebesgue" grids (thick curves, bottom two) are converging to the Lebesgue grid as $O(1/N^2)$ or faster.

7. DIFFERENCES IN THE LOCATION OF GRID POINTS

Since the expanded-Chebyshev and Chebyshev-of-the-second-kind-Lobatto grids perform very similarly to the Lebesgue-optimal grid, an obvious question is: how close are the locations of the

grid points for these sets of points to those of the Lebesgue-optimal grid? Figure 9 (bottom two curves) provides an answer: the grid points for these two grids are *very* close to the Lebesgue points.

As N increases, the average spacing between points decreases as $O(1/N)$. It is, therefore, not surprising that all five grids illustrated become closer to the Lebesgue-optimal grid as N increases. In the right panel, however, the differences in grid locations have been scaled by multiplication by a power of N (plus a constant chosen to stagger the curves in the same vertical arrangement as in the left panel). The top three curves have been scaled by multiplication by N, but all three scaled differences seem to be slowly growing with N. This implies that the differences between the Chebyshev-Lobatto, Legendre-Lobatto (VDM-optimal) and Erdős grids are decreasing more slowly than $1/N$, and therefore, are not really converging to the Lebesgue grid at all except in the sense that the point-to-point separation for all grids is contracting.

In contrast, the expanded-Chebyshev and Chebyshev-of-the-second-kind-Lobatto grids have been scaled by N^2 in the right panel. Both curves either level off (Chebyshev-of-the-second-kind-Lobatto) or decrease (expanded-Chebyshev), which implies that these grid points are converging to the Lebesgue grid as $O(1/N^2)$ or faster.

Of course, numerical evidence is not proof. However, the convergence of these "almost-Lebesgue" grids to the Lebesgue points as $N \to \infty$ is a plausible conjecture. Perhaps a mathematician will someday prove it, or show that the numerical evidence for $N \leq 20$ is misleading.

Hesthaven [15] gives a graph showing that the Lebesgue constants for the Chebyshev-of-the-second-kind-Lobatto grid become *worse* (that is, larger) than those of the Legendre-Lobatto grid for large N. However, his analytic formula for these points, given without a number label just above his Conjecture 3.3, is not correct; the grid points are given by $x_k = \cos(t_k)$ where t_k is a solution of the transcendental equation $N\cos(Nt)\sin(t) - \sin(Nt)\cos(t) = 0$, which has no simple closed form solution even in the limit $N \to \infty$ where the smallest nonzero root is $t_2 \sim \tau/N$ where $\tau \approx 4.493409$ solves $\tan(\tau) - \tau = 0$. His Conjecture 3.3 and Figure 1, are therefore, suspect.

8. GEGENBAUER HYPOTHESIS

The close agreement between the Chebyshev-of-the-second-kind-Lobatto grid and the Lebesgue-optimal grid led us to conjecture that perhaps the Lebesgue-optimal grid was the Lobatto grid for Gegenbauer polynomials of some $\alpha(N)$ where α is the exponent of the weight function in the orthogonality integral that defines the polynomial set. This hypothesis is false. However, our experiments showed that the difference between the Gegenbauer-Lobatto and Lebesgue-optimal grids was minimized by choosing

$$\alpha \approx \frac{1}{2} - \frac{0.47}{N}. \tag{24}$$

For $N < 14$, the best Gegenbauer-Lobatto grid was closer to the Lebesgue-grid than the expanded-Chebyshev grid, but for a large number of points, however, the expanded-Chebyshev grid is superior to any Gegenbauer grid, no matter how much we tweak α. The analytic form of the Lebesgue-optimal grid points still remains an enigma wrapped in a mystery.

9. INTERPOLATION ON THE TRIANGLE

Our results for the *interval* extend very easily to quadrilaterals because the usual grids on rectangular elements are tensor products of grids for the interval, and therefore, are effectively one dimensional. In contrast, the Erdős-optimal and Fekete-optimal grids for the triangle, which have been computed by Chen and Babuška [1–3], are decidely *not* tensor-product grids.

Unfortunately, there has not been a comprehensive comparison of different grids for the triangle. However, Chen and Babuška [10] (Table 2) show that the Lebesgue constants for polynomials of only degree nine on the tetrahedron in three space dimensions are already 42% larger for the Fekete-optimal grid than for the Lebesgue-optimal grid. This suggests that the differences

between grids are magnified as the dimensionality increases, but are still modest. But this is only speculation.

It is not speculative, however, to note that issues of numerical ill-conditioning will rapidly escalate with the dimensionality. On the triangle, polynomials of eighth degree require 45 grid points in two dimensions versus only nine on the interval.

Another key issue which does not arise in one dimension is mixing triangular and quadrilateral elements. The Legendre-Lobatto grid is the most popular choice for spectral elements on quadrilaterals. For the Fekete (VDM-optimal) grid on the triangle, the points on each side are those of the Legendre-Lobatto grid on the interval, which makes it easy to match triangular elements with rectangular elements. Taylor and Wingate [6], therefore, advocate a Fekete grid for the triangle. Királyfalvi and Szabó solve the same problem in a different way: by using the boundary points of the Erdős-optimal grid for the triangle [3] to define a tensor product grid for quadrilateral elements.

Another option is to set up a grid for the triangle which is Legendre-Lobatto along the sides, like the Fekete grid, but adjust the positions of the interior grid points to minimize the Lebesgue constants, like a Lebesgue-optimal grid.

Obviously, the choice of the best grid for the triangle and tetrahedron is far from settled.

10. SUMMARY

The Lebesgue-optimum grid is the best choice on the basis of the criteria investigated here because it performs well according to all three measures of optimality. The expanded-Chebyshev grid may be substituted for programming convenience because it performs almost as well as the Lebesgue grid and is available as a simple analytical formula, instead of as a table or a nonlinear programming computation [4]. The Chebyshev-of-the-second-kind-Lobatto is almost equally effective, but requires a short program, rather than a formula, such as the FORTRAN program of Appendix C of [22]. The differences between the locations of these very-good-and-almost-equivalent grids are very small as shown in Figure 9. We conjecture that these "almost-Lebesgue" grids converge to the Lebesgue grids as fast as $O(1/N^2)$ as $N \to \infty$, based on the numerical evidence.

Lebesgue-optimality seems to be the most useful criterion of optimality: the grid that is best according to this criterion is nearly Cauchy-optimal and Fekete-optimal, too, and further allows the longest stable timestep. In contrast, the Chebyshev-interior grid, which is the best for Cauchy optimality, is mediocre for the other two measures of performance. The VDM-optimum grid is similarly mediocre when measured by Cauchy-optimality or Lebesgue-optimality. VDM-optimality seems to be the least discriminating of the four accuracy criteria because,

(i) all grids are very similar in performance according to this measure, and
(ii) it is not justified by a convergence theorem [Cauchy-optimality] or a functional analysis bound [Lebesgue-optimality].

Our third theme, however, is that the accuracy differences between the various grids on the various optimality measures are so small that these differences should always take second place to convenience. These differences are always smaller than a factor of two, usually much smaller. Because the error in approximating a smooth $f(x)$ normally decreases exponentially fast with N [20–22], one can usually recover a factor of two in error merely by increasing the number of interpolation points by one. The monograph [20], therefore, argues that the factor-of-two theoretical difference between the pseudospectral algorithm and Galerkin's method is always insignificant, and accuracy should take second place to convenience. Since the differences in Lebesgue constant or Cauchy optimality here are mostly much smaller than a factor of two, this goes double here: ease-of-programming and other factors unrelated to accuracy should choose the grid, not a five percent variation in the Lebesgue constant or $||\omega(x)||_\infty$.

Our results also apply to two-dimensional interpolation on a quadrilateral since the usual grids are just tensor products of one-dimensional grids. No study like ours has yet been performed for interpolation on a triangle where the Lebesgue-optimum grid is most emphatically not a tensor-product grid. However, if we are willing to speculate that the modest differences between different grids found here carry over to the triangle, then the implication is that enthusiasm for one grid over another merely because it is Lebesgue-optimal or Erdős-optimal or Fekete-optimal is probably misplaced.

Instead, the relevant questions for the triangle are the following: what grid is most well-conditioned? What grid allows conformal matching of triangular to quadrilateral elements in an unstructured grid that includes both kinds of elements? As long as a grid on the triangle is "Chebyshev-like", that is, with a density of points that varies smoothly from low near the center to high near the edges and especially the vertices, criteria such as sensitivity to roundoff error, ease of programming, conformality with nontriangular elements, and maximum timestep are likely to be far more important in choosing the "best" grid than slight variations in accuracy.

There is a need for a study of optimality on the triangle that will not only compare many species of grids as here, but also supplement accuracy criteria with issues of triangle-quadrilateral blending and sensitivity to roundoff error.

REFERENCES

1. Q. Chen and I. Babuška, Approximate optimal points for polynomial interpolation of a real function in an interval and in a triangle, *Computer Methods in Applied Mechanics and Engineering* **128**, 405–417, (1996).
2. Q. Chen and I. Babuška, The optimal symmetrical points for polynomial interpolation of real functions in the tetrahedron, *Computer Methods in Applied Mechanics and Engineering* **137**, 89–94, (1996).
3. Q. Chen and I. Babuška, Adaptive procedure for approximating functions by continuous piecewise polynomials, *Communications in Numerical Methods in Engineering* **12**, 483–495, (1996).
4. J.R. Angelos, R.H. Kaufman, Jr., M.S. Henry and T.D. Lenker, Optimal nodes for polynomial interpolation, In *Approximation Theory VI*, (Edited by Chui, Schumaker and Ward), pp. 17–20, Academic, New York, (1989).
5. L.P. Bos, Bounding the Lebesgue functions for Lagrange interpolation in a simplex, *Journal of Approximation Theory* **38**, 43–59, (1983).
6. M.A. Taylor and B.A. Wingate, The Fekete collocation points for triangular spectral elements, *SIAM Journal of Numerical Analysis*, (submitted).
7. B.A. Wingate and M.A. Taylor, The natural function space for triangular spectral elements, *SIAM Journal of Numerical Analysis*, (submitted).
8. B.A. Wingate and M.A. Taylor, A Fekete point triangular spectral element method; Application to the shallow water equations, *SIAM Journal of Numerical Analysis*, (submitted).
9. S.J. Sherwin and G.E. Karniadakis, A new triangular and tetrahedral basis for high-order finite elements: Applications to the incompressible Navier-Stokes equations, *International Journal for Numerical Methods in Engineering* **38**, 3775, (1995).
10. S.J. Sherwin and G.E. Karniadakis, A triangular spectral element method: Applications to the incompressible Navier-Stokes equations, *Computer Methods in Applied Mechanics and Engineering* **123**, 189–229, (1995).
11. S.J. Sherwin and G.E. Karniadakis, Tetrahedral *hp* finite elements: Algorithms and flow simulations, *Journal of Computational Physics* **124**, 14–45, (1996).
12. S.J. Sherwin, Hierarchical *hp* finite elements in hybrid domains, *Finite Elements in Analysis and Design* **27**, 109–119, (1997).
13. B.A. Wingate and J.P. Boyd, Spectral element methods on triangles for geophysical fluid dynamics problems, In *Proceedings of the Third International Conference on Spectral and High Order Methods*, (Edited by A.V. Ilin and L.R. Scott), pp. 305–314, Houston, TX, (1996).
14. G. Királyfalvi and B.A. Szabó, Quasi-regional mapping for the *p*-version of the finite element method, *Finite Elements in Analysis and Design* **27**, 85–97, (1997).
15. J.S. Hesthaven, From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex, *SIAM Journal of Numerical Analysis* **35**, 655–676, (1998).
16. M. Feketem, Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, *Math. Zeits.* **17**, (1923).
17. L.Brutman, On the Lebesgue function for polynomial interpolation, *SIAM Journal of Numerical Analysis* **15**, 694–704, (1978).
18. L. Fejér, Bestimmung derjenigen Abszissen eines Intervalles für welche die Quadratsumme der Grundfunctionen der Lagrangeschen Interpolation im Intervalle $[-1, 1]$ ein möglichst kleines Maximum besitzt, *Ann. Scuola* **1**, Norm. Sup. Pisa, 3–16, (1932).

19. L. Brutman, Lebesgue functions for polynomial interpolation—A survey, *Annals of Numerical Mathematics* **4**, 111–128, (1997).
20. J.P. Boyd, *Chebyshev and Fourier Spectral Methods*, Springer-Verlag, New York, (1989).
21. B. Fornberg, *A Practical Guide to Pseudospectral Methods*, Cambridge University Press, New York, (1996).
22. C. Canuto, M.Y. Hussaini, A. Quarteroni and T.A. Zang, *Spectral Methods for Fluid Dynamics*, Springer-Verlag, New York, (1988).
23. C. Bernardi and Y. Maday, *Approximations Spectrales de Problémes aux Limites Elliptiques*, Springer-Verlag, Paris, France, (1992).
24. D. Fundaro, *Volume 1 of Lecture Notes in Computational Science and Engineering*, Springer-Verlag, Heidelberg, (1997).